Next Item Recommendation w/ Self-Attention

Shuai Zhang, Lina Yao, Yi Tay, Aixin Sun
Next Item Recommendation
(Sequence-aware)

- **Wu et al.** 2017. Recurrent Recommender Networks.
Next Item Recommendation
(Deep Learning)

- MLPs, ConvNets, Auto Encoders, RNNs, Attention.
- **Tang et al.** 2018. Personalized Top-N Sequential Recommendation via Convolutional Sequence Embedding
- **Yuyun Gong and Qi Zhang**. 2016. Hashtag Recommendation Using Attention-based Convolutional Neural Network
Attention


the agreement on the European Economic Area was signed in August 1992.
Self-Attention

AVATAR

GUARDIANS OF THE GALAXY

INTERSTELLAR

STAR WARS

ROCKY

A CLOCKWORK ORANGE

The Godfather

STAR WARS

?
Self-Attention

- Avatar
- Guardians of the Galaxy
- Interstellar
- Star Wars
Self-Attention

- Avatar
- Guardians of the Galaxy
- Interstellar
- Star Wars
Self-Attention

The FBI is chasing a criminal on the run.
The FBI is chasing a criminal on the run.
The FBI is chasing a criminal on the run.
The FBI is chasing a criminal on the run.
The FBI is chasing a criminal on the run.
The FBI is chasing a criminal on the run.
The FBI is chasing a criminal on the run.
The FBI is chasing a criminal on the run.
The FBI is chasing a criminal on the run.
The FBI is chasing a criminal on the run.
The FBI is chasing a criminal on the run.
The FBI is chasing a criminal on the run.
Self-Attention

\[ \begin{align*}
X_t^u & \rightarrow \text{Nonlinear Layer} \rightarrow \text{Identity Mapping} \\
& \rightarrow \text{Nonlinear Layer} \rightarrow \text{Scale} \\
& \rightarrow \text{Mask} \rightarrow \text{Softmax} \\
& \rightarrow \text{Attention Map} s_t^u \\
& \rightarrow a_t^u
\end{align*} \]
Full Model

Interaction History

Self-Attention Module

Long-term Preference

Item Embedding

Mean

Euclidean Distance

Item Embedding

Short-term Intent

Euclidean Distance

Movie Genre
Evaluation Metrics
Results

- Hit ratio (HR@N)
- mean reciprocal rank (MRR)
## Results

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Metric</th>
<th>POP</th>
<th>BPRMF</th>
<th>MC</th>
<th>FPMC</th>
<th>HRM</th>
<th>PRME</th>
<th>TransRec</th>
<th>Caser</th>
<th>AttRec</th>
<th>Improv.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML-100K</td>
<td>HR@50</td>
<td>0.2142</td>
<td>0.3754</td>
<td>0.4115</td>
<td>0.4783</td>
<td>0.4821</td>
<td>0.4411</td>
<td>0.4634</td>
<td>0.4667</td>
<td><strong>0.5273</strong></td>
<td>9.38%</td>
</tr>
<tr>
<td></td>
<td>MRR</td>
<td>0.0388</td>
<td>0.0616</td>
<td>0.0662</td>
<td>0.0925</td>
<td>0.0889</td>
<td>0.0837</td>
<td>0.0827</td>
<td>0.0799</td>
<td><strong>0.0981</strong></td>
<td>10.35%</td>
</tr>
<tr>
<td>ML-HetRec</td>
<td>HR@50</td>
<td>0.1065</td>
<td>0.1462</td>
<td>0.1903</td>
<td>0.2321</td>
<td>0.2380</td>
<td>0.2357</td>
<td>0.1912</td>
<td>0.2144</td>
<td><strong>0.2964</strong></td>
<td>24.54%</td>
</tr>
<tr>
<td></td>
<td>MRR</td>
<td>0.0177</td>
<td>0.0215</td>
<td>0.0359</td>
<td>0.0489</td>
<td>0.0486</td>
<td>0.0500</td>
<td>0.0337</td>
<td>0.0387</td>
<td><strong>0.0592</strong></td>
<td>18.40%</td>
</tr>
<tr>
<td>ML-1M</td>
<td>HR@50</td>
<td>0.1440</td>
<td>0.2378</td>
<td>0.3419</td>
<td>0.4209</td>
<td>0.4311</td>
<td>0.4449</td>
<td>0.3358</td>
<td>0.4811</td>
<td><strong>0.5223</strong></td>
<td>8.56%</td>
</tr>
<tr>
<td></td>
<td>MRR</td>
<td>0.0231</td>
<td>0.0368</td>
<td>0.0654</td>
<td>0.1022</td>
<td>0.0873</td>
<td>0.1044</td>
<td>0.0561</td>
<td>0.0925</td>
<td><strong>0.1172</strong></td>
<td>12.26%</td>
</tr>
<tr>
<td>Android App</td>
<td>HR@50</td>
<td>0.1194</td>
<td>0.1738</td>
<td>0.1802</td>
<td>0.1990</td>
<td>0.2001</td>
<td>0.1686</td>
<td>0.2016</td>
<td>0.1426</td>
<td><strong>0.2187</strong></td>
<td>8.48%</td>
</tr>
<tr>
<td></td>
<td>MRR</td>
<td>0.0228</td>
<td>0.0287</td>
<td>0.0355</td>
<td>0.0355</td>
<td>0.0295</td>
<td>0.0237</td>
<td>0.0306</td>
<td>0.0231</td>
<td><strong>0.0415</strong></td>
<td>16.90%</td>
</tr>
<tr>
<td>Health / Care</td>
<td>HR@50</td>
<td>0.0337</td>
<td>0.0900</td>
<td>0.0786</td>
<td>0.1128</td>
<td>0.0965</td>
<td>0.0843</td>
<td>0.0962</td>
<td>0.0768</td>
<td><strong>0.1272</strong></td>
<td>12.77%</td>
</tr>
<tr>
<td></td>
<td>MRR</td>
<td>0.0171</td>
<td>0.0188</td>
<td>0.0245</td>
<td>0.0258</td>
<td>0.0183</td>
<td>0.0119</td>
<td>0.0232</td>
<td>0.0146</td>
<td><strong>0.0277</strong></td>
<td>7.36%</td>
</tr>
<tr>
<td>Video Game</td>
<td>HR@50</td>
<td>0.0609</td>
<td>0.1630</td>
<td>0.1708</td>
<td>0.2226</td>
<td>0.2150</td>
<td>0.1855</td>
<td>0.2035</td>
<td>0.1438</td>
<td><strong>0.2414</strong></td>
<td>8.45%</td>
</tr>
<tr>
<td></td>
<td>MRR</td>
<td>0.0126</td>
<td>0.0277</td>
<td>0.0381</td>
<td>0.0451</td>
<td>0.0337</td>
<td>0.0263</td>
<td>0.0349</td>
<td>0.0248</td>
<td><strong>0.0496</strong></td>
<td>9.98%</td>
</tr>
<tr>
<td>Tools / Home</td>
<td>HR@50</td>
<td>0.0319</td>
<td>0.0559</td>
<td>0.0384</td>
<td>0.0535</td>
<td>0.0488</td>
<td>0.0465</td>
<td>0.0658</td>
<td>0.0424</td>
<td><strong>0.0775</strong></td>
<td>17.78%</td>
</tr>
<tr>
<td></td>
<td>MRR</td>
<td>0.0061</td>
<td>0.0099</td>
<td>0.0093</td>
<td>0.0129</td>
<td>0.0086</td>
<td>0.0076</td>
<td>0.0112</td>
<td>0.0071</td>
<td><strong>0.0148</strong></td>
<td>14.73%</td>
</tr>
<tr>
<td>Digital Music</td>
<td>HR@50</td>
<td>0.0436</td>
<td>0.1621</td>
<td>0.1307</td>
<td>0.1580</td>
<td>0.1998</td>
<td>0.1559</td>
<td>0.1894</td>
<td>0.1327</td>
<td><strong>0.2205</strong></td>
<td>10.36%</td>
</tr>
<tr>
<td></td>
<td>MRR</td>
<td>0.0073</td>
<td>0.0277</td>
<td>0.0320</td>
<td>0.0322</td>
<td>0.0310</td>
<td>0.0243</td>
<td>0.0300</td>
<td>0.0228</td>
<td><strong>0.0467</strong></td>
<td>45.03%</td>
</tr>
<tr>
<td>Garden</td>
<td>HR@50</td>
<td>0.0319</td>
<td>0.0965</td>
<td>0.0946</td>
<td>0.1525</td>
<td>0.1593</td>
<td>0.1573</td>
<td>0.1486</td>
<td>0.1632</td>
<td><strong>0.2177</strong></td>
<td>33.39%</td>
</tr>
<tr>
<td></td>
<td>MRR</td>
<td>0.0049</td>
<td>0.0105</td>
<td>0.0333</td>
<td>0.0408</td>
<td>0.0255</td>
<td>0.0266</td>
<td>0.0257</td>
<td>0.0277</td>
<td><strong>0.0459</strong></td>
<td>12.50%</td>
</tr>
<tr>
<td>Instant Video</td>
<td>HR@50</td>
<td>0.1240</td>
<td>0.2350</td>
<td>0.1650</td>
<td>0.2120</td>
<td>0.2430</td>
<td>0.1910</td>
<td>0.2570</td>
<td>0.1620</td>
<td><strong>0.2790</strong></td>
<td>8.56%</td>
</tr>
<tr>
<td></td>
<td>MRR</td>
<td>0.0173</td>
<td>0.0376</td>
<td>0.0426</td>
<td>0.0541</td>
<td>0.0414</td>
<td>0.0367</td>
<td>0.0441</td>
<td>0.0275</td>
<td><strong>0.0634</strong></td>
<td>17.19%</td>
</tr>
<tr>
<td>LastFM</td>
<td>HR@50</td>
<td>0.1314</td>
<td>0.3659</td>
<td>0.1682</td>
<td>0.2808</td>
<td>0.3733</td>
<td>0.2503</td>
<td>0.3785</td>
<td>0.1756</td>
<td><strong>0.3901</strong></td>
<td>3.06%</td>
</tr>
<tr>
<td></td>
<td>MRR</td>
<td>0.0224</td>
<td>0.1062</td>
<td>0.0645</td>
<td>0.0869</td>
<td>0.1209</td>
<td>0.1276</td>
<td>0.1147</td>
<td>0.0343</td>
<td><strong>0.1312</strong></td>
<td>2.82%</td>
</tr>
<tr>
<td>MovieTweetings</td>
<td>HR@50</td>
<td>0.1687</td>
<td>0.1749</td>
<td>0.3314</td>
<td>0.3417</td>
<td>0.3105</td>
<td>0.3286</td>
<td>0.2755</td>
<td>0.3332</td>
<td><strong>0.3602</strong></td>
<td>5.41%</td>
</tr>
<tr>
<td></td>
<td>MRR</td>
<td>0.0204</td>
<td>0.0231</td>
<td>0.0700</td>
<td>0.0674</td>
<td>0.0534</td>
<td>0.0476</td>
<td>0.0421</td>
<td>0.0585</td>
<td><strong>0.0811</strong></td>
<td>15.86%</td>
</tr>
</tbody>
</table>
Parameter Sensibility

(a) Movielens 100K
(b) Amazon Garden

(a) Effects of the weight $\omega$
(b) Effects of the sequence length $L$
Possible Extensions

- Attention instead of pondered mean for combination of long and short term modules.
- Attention instead of mean (transformer, etc)
- Longer \(\text{?histories?}\)
- Extension to implicit feedback datasets.